Interview with Endel Tulving

Endel Tulving is University Professor at the University of Toronto. He was born in 1927 in Estonia, and came to Canada in 1949. He did his undergraduate work at the University of Toronto and graduate work at Harvard. He has been teaching in the Department of Psychology at Toronto since 1956, with a four-year stint at Yale in the early 1970s. His research has been concerned with human memory. His discoveries and the concepts he has introduced to the field include subjective organization, input and output interference, the distinction between availability and accessibility of stored information, retrieval cue, cue-dependent forgetting, Tulving–Colotla measure of primary memory, recognition failure of recallable words and names, the reduction method of determining trace structures, Tulving–Wiseman function of the relation between recognition and cued recall, encoding specificity, synergistic ecphory, encoding/retrieval interaction, perceptual and ecphoric similarity relations in recognition, the distinction between episodic and semantic memory, the distinction between noetic and autonoetic consciousness, stochastic independence between perceptual priming and explicit memory measures, and the quasimemory system of perceptual priming.

JOCN: Let us begin at the macrolevel, human memory. What does it mean to have a theory of human memory? Or, put differently, what must a theory of human memory contain for it to reflect the known complexity of the process?

ET: I trust that you are talking about theories *about* human memory, rather than theories *of* memory as such. A theory of memory would be something like a theory of light, or a theory of evolution, that tells you what it (light, evolution) is, and how its phenomena must be what they are because of what the theory says they should be, or how it is perfectly sensible that they are what they are. There have not been any theories *of* memory of this kind, and it is a reasonably safe bet that there never will be any.

Now, theories *about* memory are concerned with selected, restricted sets of phenomena of memory; they thereby escape facing the problem of complexity of memory as such. In cognitive psychology, these "local" theories represent, as you well know, variations on the general theme of cognition as information processing. So, memory is thought of as consisting of the processes of encoding, storing, and retrieving information, and each of these processes consists of subprocesses. Phenomena of memory are explained in terms of the characteristics of these processes, and their interactions.

As to the second part of your question, a theory of something is an explanation of that something, and to explain something means different things to different people. Some think that memory will be explained—I even fear that some might claim it will be understood—when the underlying synaptic mechanisms have been identified. There are others who think that a computer program, or a mathematical model, can explain memory, or at least some of its phenomena. Between these extremes are other, potentially more fruitful, basic orienting attitudes. So, given these differences, you will get different suggestions as to what the theory must contain to cope with the known process.

JOCN: Well, we are asking you. In particular, must a cognitive theory about memory that would please you be stated in a way that could be tested by brain scientists?

ET: Sure! But an even better idea might be to demand that a cognitive theory be stated in a way that the Almighty himself could pass judgment on. The point is that any interesting cognitive theory about memory (or a cognitive theory about some interesting phenomenon of memory)—that is, *behavior* or *experience* that can be classified under the rubric—is utterly beyond the pale of most methods and techniques of today's brain science. Things may change tomorrow, of course, and then we may want to include physiological or even physical reality as a criterion for evaluating cognitive theories, but right now the insistence (you said "must it be?") on it would bring about a quick demise of cognitive theories of memory.

Look, we have great difficulties making *psychological* sense of many things we observe about memory, that is, just making up plausible, reasonably economical, and internally consistent stories about those phenomena that have caught the intellectual fancy of a particular generation of practitioners. If we had to start worrying about whether a favorite theory of ours is really true, that is, how Mother Nature planned it all, or if we began expecting that people in some other branch of the science of memory do so, we and they would probably freeze in thought instantly.

JOCN: When people commonly think of memories, of recalling past experience, they imagine the information is somehow recorded in a particular site in the brain. Is there anything wrong with that folk notion, the same notion most brain scientists believe to be true? Or jumping ahead, what are, say the five facts of memory that brain scientists ought to be considering as they pursue the physical dimensions of this problem?

ET: Let me just answer the first part of the question. We will see later whether we get around to the second part.

There is nothing wrong *in principle* with the idea that the information that is necessary for remembering something is recorded in a particular site in the brain. It is

almost a tautology. When an event occurs that a person perceives and subsequently remembers, some changes must occur in the brain. That is, the brain is different before and after that event, or experience, or whatever. We can call that before–after difference the engram, or the memory trace, or the representation of the event, or whatever. The engram, by definition, must be localizable somewhere in the brain. This is why the idea of the engram was scientifically respectable long before Lashley began his famous, and eventually inconclusive, search for it, and why it is still so.

JOCN: Why did you say *in principle* nothing is wrong with the idea of physical storage?

ET: I said "nothing wrong in principle," because in practice there does exist a problem, quite apart from the complexity of it all. The concept of engram, however it is labeled, has mesmerized many brain scientists into acting as if there was nothing more to the problem of memory and the brain than the engram and its characteristics, including its location in the overall structure. Many of these brain scientists, like many common people, do not seem to realize that the engram is an unfinished thought about memory, that it is at best only one-half the story of memory. This being so, when they concentrate on the one-half and ignore the equally essential other half, they may be doing the right thing, or they may be doing the wrong thing. It is difficult to tell in advance. I wish someone would tell me whether they are unaware of the other half, whether they are preoccupied with the engram out of sheer inertia, or whether they have reflected deeply on the matter and deliberately decided that identification of storage sites of memories is the number one priority, for such and such compelling reasons.

JOCN: By the "other half" you presumably mean retrieval? Storage of information, or engram, is one-half, you say, and retrieval is the other half?

ET: Exactly, although I hope you realize that talking about two "halves" means simplifying the matter greatly. Also, talking about the two "halves" misses the most important feature of each, namely that neither can work separately, that is, that storage and retrieval processes receive their "identity" from the interaction between them.

A biological memory system differs from a mere physical information-storage device by virtue of the system's inherent capability of using the information in the service of its own survival, which, we are told, is any biological system's first priority. The Library of the Congress, a piece of videotape, or a Cray supercomputer, and many other devices that store information, could not care less about their own survival. So, anyone who is interested in memory, but looks only at the storage side of things, is essentially ignoring the fundamental distinction between dead and living storage systems, that is, ignoring the essence of *biological* memory.

JOCN: What do cognitive psychologists know about storage and retrieval that makes the distinction so crucial as you seem to be implying?

ET: Cognitive psychologists "discovered" retrieval and figured out how to separate it analytically and experimentally from storage in the 1960s. The predecessor of memory research in experimental psychology was "verbal learning," the study of learning and retention of verbal materials. Its pretheoretical thinking was greatly inspired by conditioning. It, too, suffered from the preoccupation with storage, although the concept was then labeled "association." The interesting thing is that students of verbal learning were unaware of their "storage bias," for the simple reason that the distinction between storage and retrieval had not yet been made: You cannot be aware of something that does not exist. I should mention parenthetically that for the first six years or so of my post-Ph.D. life I too was one of those happy verbal learners who did not lose any sleep over the storage–retrieval distinction. When things changed, the field of verbal learning essentially died.

It would not be difficult to argue that the "discovery" of retrieval *processes* permanently revolutionized the field of memory research in cognitive psychology. And yet, as I see things, that revolution has not yet reached brain scientists. I have seen little evidence that retrieval processes occupy their thoughts or shape their activities.

JOCN: Brain scientists ignore the larger question because they do not have any idea how to study the issue from a "neural systems" point of view. The synapse is where the light is shining, which is to say it is something that can be studied. Also, at a superficial level, it makes sense that information storage ought to reflect structural and physiological changes at the synapse. Yet, no one thinks memories are stored at synapse, that is, if you push them on it. So, that suggests brain scientists would be well served to have the problem properly characterized. Any comments?

ET: I hope you are not implying that you or I should tell the brain scientists what their proper problems are! Such an action would generate only heat, and absolutely no light. Surely they are doing the best they can. Besides, there is nothing wrong with studying what is possible to study, and not worrying about what is not; the same venerable principle holds in all sciences. And what is possible to study at a given time depends to a large extent on the amount of relevant knowledge that has already been accumulated.

Having said that, let me nevertheless offer a possibly pertinent thought.

Of course memories are not stored at synapses. But I think it is useful to contemplate the possibility that they are *not stored anywhere else in the brain either*. The whole issue of where or, more important, how memories are stored in the brain may turn out to be an incorrect formulation of the problem, despite its seductively enticing allure. And the source of such incorrect formulation may lie in the single-minded preoccupation with the

storage, or the engram, and sometimes even *identification* of storage with memory. This preoccupation with the physical changes that follow from an experience that can be remembered seems to be accompanied by a rather conspicuous neglect of retrieval processes.

JOCN: So, you are not telling brain scientists what their proper problems are, but you are telling them where they seem to be missing the boat?

ET: I am just telling *you* about my casual, and possibly quite incorrect, impressions about the brain science side of things in memory research. If a brain scientist's curiosity is piqued about this "vague" talk about the crucial distinction between storage and retrieval, and if he then decides to do something about it, it is his right and privilege. I would not want to prescribe research problems for people in a different discipline any more than I would care about such prescriptions coming from them.

But I cannot deny that I would be pleased to see the whole issue openly discussed, and at least some brain scientists experimenting with formulations that place the emphasis on the nature of brain *activity* that subserves or produces the kind of mental *activity* that is identified with memory in cognitive psychology.

JOCN: And synaptic mechanisms are not part of that activity? Surely those mechanisms are involved, and perhaps even determine, storage?

ET: *The key process of memory is retrieval.* The storage or engram alone, in the absence of retrieval, is no better than no storage and no engram at all. If you know something, or if you have stored information about an event from the distant past, and never use that information, never think of it, your brain is functionally equivalent to that of an otherwise identical brain that does not "contain" that information. And it is but a small step from this idea of functional equivalence to the idea of structural equivalence: An engram does *not exist independently of retrieval*, that is, a brain "containing" a nonretrieved engram is structurally equivalent to an otherwise identical brain that does not "contain" that particular engram.

JOCN: Now you are losing me. Earlier you said the engram exists, it is physical, and it is localized, and now you are saying it does not exist? Is this a new version of one of those famous Zeno's paradoxes?

ET: It may sound like a paradox, but it is not. A physical engram exists, but it cannot be identified *as an engram* by any physical means, it can be identified only through its biological–psychological action, through the retrieval *process*.

As a scientist I am compelled to the conclusion—not postulation, not assumption, but conclusion—that there must exist certain physical–chemical changes in the nervous tissue that correspond to the storage of information, or to the engram, changes that constitute one of the necessary conditions of remembering. (The alternative stance, that it may be possible for any behavior or any thought to occur independently of physical changes in the nervous system, as all your good readers know, is sheer mysticism.) However, if the engram is a kind of an entity that manifests itself only in activity, or retrieval, then we might conjecture that the physical changes resulting from an experience *do not exist as an engram in the absence of that activity.* And we can also imagine that the engram, *qua engram*, is not detectable in its quiescent state, that is, in the absence of retrieval, *with any physical technique*.

Let me give you a very simple, and in many ways inadequate, analogy that illustrates, or at least hints at, the kind of a situation we have on our hands.

Wind is the movement of air molecules. To create something that we identify as wind, two necessary conditions must be satisfied: (1) the presence of the air molecules in sufficient quantity and density, and (2) the operation of some source of energy that sets the molecules in more-or-less coordinated motion of sufficient velocity. Now think of an analogy with memory: (1) the blowing wind is the brain activity that subserves the experience of remembering, (2) the air molecules constitute the physical substrate of the activity, the engram, and (3) the energizing force is the retrieval cue that "activates" or "ecphorizes" the engram. Thus, wind is particles and energy; remembering is engram and retrieval. In this analogy, the relation between the experience of remembering and the engram is the same as the relation between the wind and air molecules: the second member of each pair is a necessary condition of the first.

Now, what would you think of a sage who decides to identify the physical substrate, the "engram," of the wind, and starts searching for it, assuming it to be a special entity of some kind, an entity different from other like entities that do not produce wind? Since no specifically identifiable "engram" in fact exists—the air molecules that can be set in motion do not differ from those that never are set in motion—our sage is going to be spending a lot of time on his chosen problem, unaware that the problem is created by him himself, by his initial presupposition, rather than by nature.

The brain scientist who is looking for a *special change* at the synapse, one that results from some experience, one that represents or stands for memory, and one that *is different from* the synaptic activity that subserves other kinds of behavior or cognition, or other kinds of physiological activity altogether, takes the risk of being terribly frustrated because, by looking for something that in fact does not exist independently of something else, he is doomed to failure, regardless of how hard he tries and how long he persists. He will have become a victim of the second of the two kinds of obstacles that lie in the path of all explorers of nature—those placed there by nature and those placed there by man.

JOCN: You may have just captured a brain scientist's worst nightmare. But are you simply thinking of alternatives here or is your view motivated by the consideration of some body of data? One can certainly view much

of the neurobiology of synaptic change as of interest such as the work of LTP (long-term potentiation) and the fierce debate over whether or not such phenomena are pre- or postsynaptic in nature. Yet one can also observe whatever it is, there is no demonstration it has anything to do with psychological memory. Is it this kind of thing that concerns you and finds you playing with alternative formulations of the problem or is your view driven from psychological data? Alternatively, is it driven by a hunch on the nature of the problem?

ET: Remember, I am a cognitive psychologist. I deal with cognitive data, revealed through behavior. I cannot get as thrilled about LTP, or the issue of presynaptic versus postsynaptic protein synthesis as do neurobiologists, for the same reason that they do not get as thrilled about the intricacies of presemantic perceptual priming as I do: We lack the requisite background knowledge to appreciate each other's excitements. A friend of mine who is closer to the synapses than I am tells me, however, that the eventual outcome of the battle of pre- versus postsynaptic processes will have important practical consequences for problems such as drug addiction and its treatment. I believe him.

But you asked about the speculations I have just shared with you. Is it a hunch, an expression of a desire for an alternative, or is it suggested by data? I would say, "all of the above." It is an alternative, *suggested* by the outcomes of many cognitive psychology experiments on recall and recognition.

The experiments that are particularly relevant to the issue have been done under the general rubric of encoding specificity, or encoding/retrieval interactions. In these experiments, the identity of the to-be-remembered items is held constant, encoding conditions are manipulated (thereby creating different engrams of physically identical items), and then these engrams are probed under different retrieval conditions. Actually, only the retrieval component of this paradigm is necessary to make the point I am making; systematically varying encoding conditions would just add some bells and whistles to the tale.

Since what I have just said may not make total sense to all of your readers, let me try to illustrate it with a simple, concrete example. Imagine that you are the subject in one of my experiments, and that you see a *pair* of words, say "lady" and QUEEN. I am telling you to make sure that you remember having seen the word QUEEN in the experiment, and that I am going to test you for it. (I should mention parenthetically that the pair of words is usually presented as a part of a larger collection, but for the purposes of our story that fact is irrelevant. I should also mention that the particular words that I am testing you with, of course, are also irrelevant. Indeed the to-be-remembered items need not be single words. The item you are asked to remember might be a unique name of a well known character, such as GEORGE WASHINGTON or FLORENCE NIGHTINGALE. The outcome of the experiment is the same.)

Anyhow, after a while I ask you: "Did you see the word QUEEN in the list that you studied?" With a certain probability you say "yes," and with a certain probability you say "no." If you say "yes," I know that, because the retrieval cue that I gave you was an effective probe and the engram "responded" to it, the engram created at study must have existed at test. If you say "no," I cannot say anything very much about the engram. It may have not been created at all, it may have been "lost" while other things were going on, or it may still exist but the retrieval cue I provided happened to be an inappropriate or "wrong" probe. So, to clarify the uncertain situation I ask you another question: "What word went with the word "lady" in the list that you studied?" And, with a certain probability you now say, "QUEEN!" Thus, you cannot *recognize* (identify as previously seen in the experiment) the word you were supposed to study and to remember, and yet you can *recall* (produce) it to another cue.

What these and many other similar kinds of data tell us is that, in a fixed encoding situation that has produced a fixed engram of a particular event (such as seeing two familiar words in a particular place at a particular time), *whether the engram responds to the probe depends on the probe.* One and the same engram responds to some probes and not others. In our example the interesting thing is that it does *not* respond to the probe that is most like the specified target item, but *does* respond to a related item.

JOCN: Hold it for a second! How is it possible for me to recall something that I do not recognize? Or, in terms of your engram story, why should the engram of the word I have seen and am now trying to remember fail me when I try to retrieve it with a virtual copy of the original, and yet "respond" satisfactorily if I can go after the stored information with a cue different from the item I am trying to recover?

ET: Well, until about 20 years ago everybody (including the proverbial man in the street, the brain scientist, and the cognitive psychologist) *knew* that such a happening (recall but no recognition) was indeed not possible. Yet today it is not only possible—first year psychology students who learn about it find it perfectly reasonable when they are told what is going on—we also have a pretty good idea what is happening in the situation that I just described.

This particular outcome does not happen all the time, as I said, but only with a certain probability. In fact, the conditional probability that a studied item that can be recalled to a related cue cannot be recognized when presented by itself varies systematically from very low (near zero) to very high (unity). We know a fair amount about such systematic variability. But that is another story.

JOCN: You do not think this other story is relevant here?

ET: Not really, and certainly not directly. So, let me

return to the implications of the fact that one and the same engram responds to (or "recognizes" if you wish) some probes but not others. In our experiment, we used the most natural kind of probe possible: another complex stimulus input into the system. Johannes Müller would have said, the "most adequate stimulus." The effectiveness of this kind of probe, or cue, has been shaped by evolution over the eons. And we find that the engram is highly specific: it is "identified," "located," or "activated" (I sometimes use the term "ecphorized") by one likely looking *natural* cue and not another.

Now, the point I want to make in this connection is this: If a given engram cannot be "identified" by *some* of the most natural (i.e., biological) probes, although it can by others, how reasonable is it to expect that you could identify it using some artificial (i.e., physical) probe or detection device? What would you be looking for, and how would you know what it is that you have identified, in the absence of retrieval? Even if you could somehow identify the total pattern of physical/chemical *aftereffects* of an experienced event, in all of its intricate and elaborate detail and full-blown complexity, you would have no way of knowing or predicting what kind of a "memory" (in the sense of experience) that engram is going to produce: *that* depends on the retrieval process, and that process has not yet occurred. Aftereffects of a stimulus event do not constitute an engram. The engram consists of those components of the aftereffects that are ecphorized in the process of retrieval. This is why I suggest that it might be useful to contemplate the possibility that *the engram does not exist as an identifiable entity in the absence of retrieval, although it exists as a physically unidentifiable component of the aftereffect of the stimulus event and as a necessary condition of the biological–psychological act of remembering.*

JOCN: So what does it mean to study the synaptic mechanism of memory? How can the brain scientist distinguish between what you call aftereffects of an event and the engram, that is, the *specific* aftereffects that are "ecphorizable" in retrieval and thereby, but you say only thereby, determine whether and what the person remembers of the event?

ET: This is the sixty-four-dollar question. I suggest you go and ask a brain scientist who studies synaptic mechanisms of memory. All I am doing is suggesting that there may be a rather basic problem here, a kind of an unwarranted pretheoretical assumption, and that the problem does seem to require explicit thought. The brain scientist may be able to identify the aftereffects of an event, and he may be even able to tell the difference between the aftereffects of Event A and Event B. If he studies the synaptic mechanisms that define, or are involved in, these aftereffects for their own sake, and does not worry what those aftereffects are good for, or whether they have anything to do with specific "memories," that is, if he is willing to concede the possibility that he may be studying something other than a component of memory, then there is no problem. The problem arises when the aftereffects are called memory.

I am arguing, at the level of behavior and cognition, and on the basis of observed facts, that the engram does not exist, as a component of memory, independently of retrieval. And I have a problem in that I cannot think of any reason why the same elementary proposition does not hold equally well at the level of physical happenings in the brain. Remember that when I talk about behavior and cognition, I am also talking about the brain. The mind is only an expression of the brain, at a different level, but nevertheless an expression. Whatever the mind can do, the brain can do better; and whatever the mind does do, the brain must have done, too, in its own way.

JOCN: But could it be the case you are being too generous about what is being studied within brain science? Accepting the general concept of storage is one thing. Suggesting, however, that examining how synapses might change to reflect storage mechanisms is quite another. That too is a strong claim for the brain scientist to make given what everyone assumes to be the case, namely that storage is somehow a distributed process. At another level, however, the storage metaphor suggest specific memories ought to be lost with brain damage, in the storage areas. Is that how you would characterize what happens following brain damage?

ET: It is not only useful but important to distinguish between the storage metaphor as such, on the one hand, and the idea of the physical indeterminacy of the engram, on the other. The concept of storage is a logical necessity, even if its particular formulation and the terminology one uses are necessarily quite flexible. The data from the studies of brain damage are very clear, too, in suggesting that engrams of particular *kinds* of information at least are localized in the brain, even if the localization involves distributed information. Particular lesions do produce particular deficits in memory. It is not always clear that the deficits are caused by the damage to the areas in which the information is stored—an obvious alternative hypothesis is one of disconnection between the areas concerned with storage and those concerned with retrieval—but there is little question about the *specificity* of the loss, at some level.

All this is reasonably clear. Now, to get from these data to the idea that even particular facts that we know, or particular events that we recollect, have distinct engrams, requires a bit of extrapolation, although nothing in excess of the kind that is normal in scientific thinking. The problem is that the lesions that neuropsychologists and other cognitive neuroscientists have dealt with so far have almost invariably been massive. We can well imagine that one day in the future it is possible to produce a highly circumscribed lesion, perhaps a reversible one, that has a single consequence, say, of the person not knowing any more what a strawberry is, without any

other effects whatsoever. If so, we will have obtained strong evidence of what we now can only assume, namely that engrams even of single concepts or experiences exist, that they are real (physical), that they are specific, localized, and that they represent an essential component of the memory process. But these facts would not change the basic argument I have offered here: The "strawberry engram" will have been identified by inference from the observed failure of *retrieval*.

JOCN: So you are saying, or at least implying, that this apparent paradox between biological determinacy and physical indeterminacy of engrams comes about because of a basic conflict, or incompatibility, between biological and physical approaches? That while the engram is all those things you said—real, specific, localizable, and so on—by biological criteria, or in terms of the biological procedures (and I'm willing to lump together biology and psychology for the present purposes, as you have been doing), it is none of those things by physical criteria, in terms of physical procedures?

ET: Precisely. This seems to be in the basic nature of things. Although I would not characterize the relation between the two kinds of approach as one of conflict or incompatibility. It is rather one of dealing with *different aspects*, or different facets, of one and the same thing. And there is nothing wrong with that, of course. Indeed, it is the only way that biological science has progressed: through examining and trying to understand objects and happenings in the world from different perspectives. There need be no conflict, as long as we understand that brain scientists who study memory by studying (physical) changes at the synapse, and, say, cognitive psychologists who study memory as "synergistic ecphory" (as a joint product of storage and retrieval *processes*), have a common object of interest, and that they are simply focusing on different aspects or facets of that common object. The approach of cognitive psychology and the approach of neurobiology are *complementary*, and there is no problem whatever. The problem arises only if one assumes that the physical approach is the only one, or the most essential one, or the fundamental one, that is, the old die-hard reductionist position. Remembering is a complex, emergent, biological–psychological process of the brain. Our little chat has turned out to be an examination of the implications of such a conceptualization of memory for the study of its physical basis.

JOCN: Well, we are almost done. Any final, parting words you want to leave with the readers of the *Journal*?

ET: Yes. Students of memory of the scientific world, unite in the study of the myriad aspects of the essence of biological memory, unite in the study of the interaction between the processes of storage and retrieval!